Optimize Customer Interactions—and Profits—with Advanced Data Mining Techniques

By Usama Fayyad, Ph.D.
President & Chief Executive Officer, digiMine

Initial Web site data-mining efforts focused primarily on site statistics and transaction logs, analyzing such factors as how many hits a site received, which pages visitors viewed, how long they stayed at the site, what they purchased, and how much it cost. Such Web site analytics are fine as far as they go, but they don't go nearly far enough to help retailers get the most out of their customer interactions.

To obtain a truly comprehensive view of how your customers interact with your store, you need to be able to combine online data with data from other sources, such as demographic information and records of in-store and catalog purchases. You also must be able to segment your customers according to a variety of criteria and then analyze the specific behaviors of each segment. Finally, you need an efficient means of turning these insights into action—for example, creating promotions, campaigns, and "related items" recommendations that target particular groups of customers. With such a comprehensive set of capabilities, you can realize the ultimate benefit of data mining: gaining in-depth customer insights and acting on them to increase customers' purchases and your profits.

This article begins by looking at the different types of customer data that can be gathered and the various stages that retailers go through as they learn how to take better advantage of data-mining technology. It then shows how you can harness advanced analytics to optimize your customer interactions and improve your bottom line. Finally, it discusses the requirements needed to take full advantage of the insights gained from data mining, and explains why outsourcing may be the most cost-effective approach.

What Type of Customer Data Do You Need?

Before getting into different data-mining approaches, you need to decide, first of all, how deeply you want to understand your customers. You then need to make sure that the approach you select is capable of handling customer data that will deliver that level of understanding. Three examples of customer data you might want to gather include:

- **Demographic data**. Demographic data, such as age, geographical location, and income, has been used by direct marketers for many years to target specific groups of customers, so that promotions and campaigns can be aimed at their particular interests.

- **Transaction data**. Transaction data, typically divided into RFM segments (Recency, Frequency, and Monetary value), goes beyond general information about your customers to provide concrete data about their purchasing behaviors—information that can be helpful in predicting future purchases and targeting promotions and campaigns more effectively.

- **Clickstream data**. Clickstream data—available not only for online interactions, but also for those that take place through wireless devices, gaming systems, or cable TV—goes further than transaction data to actually capture the decision process and navigational steps that the customer goes through. Such data records every page the customer looked at leading up to the decision to either purchase or not purchase the product or service, so you know not only what they purchased (or didn't purchase), but how they arrived at that decision.

By better understanding the process leading up to the buy/no buy decision, you can more effectively influence future purchase decisions. For instance, suppose you are featuring a particular product on your home page—but when customers click through to look at it, they end up buying not that product but a different product that they link to from the featured product's page. That would tell you that you're featuring the wrong product on your home page.

Equally important, clickstream data tells you who didn't buy your products, and why. For example, if clickstream data shows that many customers who left the site without making a purchase were searching for shipping information just before they dropped out, you can conclude that the inability to find this information is what caused them to drop out. You can then test this theory by making this information readily available and seeing if that results in higher sales.

Figure 1 illustrates the differences among these three data types and shows why clickstream data delivers the richest understanding of your customers' behaviors. Of course, the best data-mining solutions take advantage of all three types of data, allowing you to correlate demographics and transaction behavior with more detailed information about the processes customers go through before making—or not making—a purchase.

[Insert Figure 1 here]

Figure 1. While demographic and transaction data provide helpful information about your customers, only clickstream data shows you the process your customers go through as they make their buy/no buy decision.

How Retailers' Use of Data Mining Is Evolving

While most online retailers today gather some sort of statistics on the efficacy of their Web site, the majority have not begun to tap the full potential of data mining. Figure 2 shows the three stages that retailers typically progress through as they come to understand how in-depth mining of customer interaction data can help them both more effectively meet their customers' needs and increase their profits.

[Insert Figure 2 here]

Figure 2. As retailers learn to make better use of the advanced data-mining technology available today, they progress from Web analytics to customer analytics to optimization of customer interactions, gaining more benefits at each stage.

Stage 1, Web Analytics, is where the late adopters are today. This stage consists of gathering Web site statistics that track customers' online behaviors—how many hits your site is getting, how many pages are being viewed, the dollar volume of sales, and so forth. This type of feedback can be helpful both for fine-tuning your Web site to better meet your customers' needs and for identifying such factors as which of your products and services generate the highest (or lowest) online revenues.

The problem, however, is that Web site statistics analyze only one aspect of your customers' interactions with you: their online behaviors. They don't capture transactions made through catalog sales or bricks-and-mortar stores; they don't include customer demographics; and they don't provide any way to segment customers in order to target specific groups more effectively. So, while Web site statistics are a good start, they only begin to scratch the surface of the benefits that advanced data-mining techniques can bring to your business.

Stage 2, Customer Analytics, adds much more depth to your understanding of your customers' interactions with your company. This stage, which is now becoming mainstream, gathers data from multiple sources, including Web site interactions, transaction data from offline purchases, and demographic data from registration forms that customers fill out. A good customer analytics solution bases its analyses not on subsets or high-level aggregations of the data, but on every individual transaction. The result is both higher accuracy and the ability to drill down to more detailed views of your customers' interactions. The richness of the data available from customer analytics gives you a much more holistic view of your customers, providing deeper insights into their behaviors, likes, and dislikes.

Good customer analytics solutions also include the ability to segment customers according to a variety of criteria and export the results to other programs, such as Excel, CRM (customer relationship management) software, or marketing campaign programs. That way, as you learn more about the behavior of a particular subset of your customers—say, high-volume purchasers in the 30-to-50 age group—you can target that group with specific marketing campaigns. Such targeted campaigns tend to produce much better results than more general approaches and are also less likely to annoy customers who aren't interested in your offer.

Stage 3, Optimization, adopted so far only by the visionaries among retailers, is the most advanced stage of data mining usage and offers the biggest payoff potential. In this stage, sophisticated data-mining algorithms sift through vast volumes of data looking for patterns that may be too subtle for humans to distinguish—and then automatically apply the discovered insights to optimize customer interactions. In other words, by tailoring recommendations and promotions to the preferences of specific groups of customers, these systems can actually change customer behaviors—whether by upselling them to a higher-priced product, cross-selling to additional related products, or even downselling to a lower-priced product in cases where the customer is abandoning a potential purchase because its price is too high. These recommendations based on data patterns produce immediate payoffs in the form of increased sales.

Recommendations can be made at three different levels: for all customers, for specific segments of customers, or on a one-to-one basis reflecting the preferences of individual customers. An example of the first type would be discovering a connection between customers who buy backpacks and those that buy jeans, and then having the Web site display a jeans promotion or link whenever customers place backpacks in their shopping carts. An example of the second type would be discovering that customers who buy a particular type of backpack prefer a specific style of jeans—and then have the link show that exact style. An example of the third type would be knowing that a particular customer prefers Diesel jeans and therefore showcasing Diesel jeans whenever you display a jeans link to that customer.

For retailers that sell through multiple channels (for instance, online, catalog, and bricks and mortar), recommendations can also take place across channels. That's important, since multi-channel customers are known to have a higher lifetime value than single-channel customers—and by marketing to them more effectively, you can further increase their value. For example, some multi-channel customers may prefer to use your Web site to view your products but then go to one of your bricks-and-mortar stores to purchase the product (say, to verify that an item of clothing fits). When you become aware of this pattern through multi-channel data analysis, you can market to these customers more effectively—for example, by emailing them a coupon that they can redeem in one of your bricks-and-mortar stores. Or, for customers who buy from you through both

online and offline channels, you can leverage your knowledge of their offline purchases to offer them online promotions tailored to their interests.

Optimization Analytics Pays Immediate Dividends

As Figure 3 illustrates, the optimization step is critical for ensuring that the insights discovered by the data-mining algorithms are promptly turned into action—and higher profits. Without this step, a weak link exists between noticing the data patterns and making corresponding changes to the Web site. Marketing managers first have to identify a pattern (which, as pointed out above, may not be all that obvious, given the vast amount of data being processed) and then have to meet with the Web site and product people to create a response. Weeks, or even months, may pass before the insight is turned into action. With optimization analytics, in contrast, the link between data and action is strong and immediate. Recommendations and promotions can be recomputed weekly—or even daily.

[Insert Figure 3 here]

Figure 3. Optimization analytics strengthen the link between identifying data patterns and acting on them to increase sales.

At this point, you may be thinking, "But isn't it risky to turn Web site control over to software?" The answer is no—because you can control what the software does and does not do through the application of business rules. For example, you might specify that purchasers of Kenneth Cole items can be linked only to other Kenneth Cole items—or that purchasers of jewelry can be linked only to other jewelry (to keep the focus on higher-priced items). Optimization analytics can also provide you with regular reports that show you the effectiveness of the systems' automatic recommendations.

J. Crew, a $75- million retailer of men's and women's apparel, shoes, and accessories, is one company that can testify to the success of optimization analytics. They had previously used a cumbersome manual procedure to recommend similar and complementary styles to online purchasers. In the fall of 2002, the company moved up to optimization analytics. They found that automatically generated recommendations generated twice as many sales as the manually generated ones. Over the course of a year, that translated into a $7 million increase in sales.

Nordstrom.com—the online subsidiary of Nordstrom, Inc., one of the nation's leading fashion specialty retailers—also recently moved up to optimization analytics. One goal was to improve the effectiveness of its "Related Items" merchandising, which suggests items to shoppers as they browse. After setting up an optimization analytics system, Nordstrom found that the solution paid for itself in the first few months. For example, in January 2002, the automation of the Related Items merchandising increased monthly sales revenue by 41 percent over manually generated recommendations, while also saving the merchandising team more than 100 hours a month.

Not a Job for Amateurs

As the foregoing examples illustrate, advanced data mining can produce dramatic results—but taking full advantage of its capabilities is not a simple matter. In fact, a full 85 percent of in-house data-mining projects ultimately fail—and 40 percent don't even get off the ground.

To set up an effective data-mining system, you need three components:

- **Data.** To say that data mining requires data seems like a statement of the obvious, but the reality of the industry is that the necessary data often is not there. Data warehouses are difficult to build—and even if you succeed in building them, they can be difficult to maintain as your business changes. For example, one company put a huge amount of effort into building a data warehouse that integrated clickstream data with online and offline transaction data. Then they changed to a different e-commerce platform, which in turn changed the way data was represented. It took them eight months to update their data warehouse to handle the new data representations. Meanwhile, they had no data on their customer interactions.

- **Algorithms.** A data mining system is only as good as its algorithms—and creating the sophisticated algorithms that make up an effective data-mining system is no easy task. In particular, when dealing with customer behavioral data, which can encompass 50 dimensions or more, you need algorithms that are capable of dealing effectively with highly dimensional data.

- **Skill sets.** Finally, setting up and maintaining an effective data-mining system requires not only a heavy resource investment, but also several distinct skill sets. You need people who have expertise in data warehousing and databases; statisticians who know how to create algorithms, deal with OLAP (online analytical processing) cubes, and choose the right algorithms for each task; and business users who understand how to accomplish business goals (such as increasing customer loyalty and average order size). Furthermore, each of these groups needs to be able to understand and communicate with the others. Most companies simply don't have all the necessary skills available in house.

These complexities are some of the reasons for the growing trend of moving extensive data-mining projects to an outsourced model. With outsourcing, you avoid taking on the risks and difficulties involved in setting up a major data-mining project while realizing all the benefits that advanced data-mining techniques have to offer. In addition, instead of having to make a large capital expenditure up front, you have much more affordable monthly payments—payments that are more than offset by the increased sales revenues that a good data-mining system will generate.

Move Up to Advanced Data Mining

To take full advantage of the promise of data mining, you need to move beyond site analytics to an in-depth, comprehensive view of your customer base—both as a whole and for the various channels and segments that make it up. You also need an automated means of turning these insights into action. Advanced data-mining techniques available today can accomplish all of the above at a very reasonable cost—particularly if you outsource the service. And the results can make a dramatic difference in your bottom line.